

Approximating welfare in large efficient markets

Ellen V. Muir* and Konstantin Borovkov†

Abstract

We consider the efficient outcome of a canonical economic market model which involves N buyers with unit demand and i.i.d. random valuations, and M sellers with unit supply whose costs are also i.i.d. random variables independent of the valuations. We approximate the joint distribution of the quantity K_α of units traded and the gains W_α from trade when there is a large number of market participants, i.e. both components of $\alpha := (N, M)$ tend to infinity. The problem is reduced to studying a process expressed in terms of two independent empirical quantile processes which, in large markets, can be approximated by appropriately weighted independent Brownian bridges. That allows us to approximate the distribution of (K_α, W_α) by that of a functional of a Gaussian process. Moreover, we give upper bounds for the approximation rate.

Keywords: Efficient market outcomes, Bayesian mechanism design, Walrasian quantity, Walrasian welfare, empirical quantile processes, Brownian bridge

2010 Mathematics Subject Classification: Primary 60F15

Secondary 91B26; 60F17; 60G15

1 Introduction and main results

Consider a canonical economic model for a market in which units of a homogeneous, indivisible good are traded among N buyers $\{i_1, \dots, i_N\}$ and M sellers $\{j_1, \dots, j_M\}$. Each buyer is interested in purchasing one unit of the good, and each seller has the capacity to produce and sell one unit of the good. Buyers are willing to purchase the good at a price not exceeding their respective reservation valuations V_1, \dots, V_N that are assumed to be independent and identically distributed (i.i.d.) random variables with a common distribution function (d.f.) F . Sellers are willing to produce at a price that is not less than their respective production costs C_1, \dots, C_M , which are

*The corresponding author. School of Mathematics and Statistics, The University of Melbourne, Parkville VIC 3010, Australia; e-mail: evmuir@unimelb.edu.au.

†School of Mathematics and Statistics, The University of Melbourne, Parkville VIC 3010, Australia; e-mail: borovkov@unimelb.edu.au.

also assumed to be i.i.d. random variables, with a d.f. G . We assume that buyers valuations and sellers costs are independent of each other and call the $(N + M)$ -tuple

$$\mathcal{R}_{(N,M)} := (V_1, \dots, V_N; C_1, \dots, C_M)$$

a realization of the unitary market $\mathcal{M} := \langle N, M, F, G \rangle$.

Such markets have been studied extensively in the literature on Bayesian mechanism design, see [9], [16] and [15]. Introductions to mechanism design can be found in Ch. 5 of [10] or in [2].

Within a market, the welfare generated by trade is defined as the sum of trading buyer valuations less the sum of trading seller costs:

$$\sum_{i \in \mathcal{N}} V_i - \sum_{j \in \mathcal{M}} C_j,$$

where $\mathcal{N} \subset \{i_1, \dots, i_N\}$ and $\mathcal{M} \subset \{j_1, \dots, j_M\}$ are, respectively, the subsets of buyers and sellers who trade in the market, so that $|\mathcal{N}| = |\mathcal{M}|$. Market welfare can be thought of as the total gain from trade experienced by market participants.

A market is said to be *efficient* if the level of market welfare is always maximised for given buyer valuations and seller costs. The quantity traded in an efficient market is known in economics as the efficient quantity, the welfare-maximising quantity, or the *Walrasian quantity*. To compute it, order buyer valuations V_i from highest to lowest: $V_{[1]} \geq \dots \geq V_{[N]}$, and order seller costs C_j from lowest to highest: $C_{(1)} \leq \dots \leq C_{(M)}$. The Walrasian quantity is then given by

$$K := \arg \max_{0 \leq k \leq N \wedge M} \sum_{i=1}^k (V_{[i]} - C_{(i)}) \equiv |\{k \in \{1, \dots, N \wedge M\} : V_{[k]} - C_{(k)} \geq 0\}|. \quad (1)$$

The *Walrasian welfare*

$$W := \sum_{i=1}^K (V_{[i]} - C_{(i)}) \quad (2)$$

is the level of welfare generated by an efficient market. Figure 1 provides an illustration for the quantities K and W .

Note that W is given by a dependent random sum of random variables and there seems to be no simple way to compute its distribution. A closed form expression for the distribution of W was given in [15], but it is cumbersome and unsuitable for further analysis. Calculating the joint distribution of (K, W) is an even more difficult problem which is, nonetheless, of significant applied interest, as knowing that distribution (or at least the covariance of K and W) is needed to determine the behaviour of quantities such as revenue (see e.g. [14]).

However, in many applications in economics one also assumes that the market is large, i.e. both N and M tend to infinity, and so determining the asymptotic distribution of (K, W) in such a situation is of substantial interest. In much of

the previous mechanism design literature, only the leading term behaviour of the mechanism outcomes was determined (which is a law of large numbers type result; e.g. it was shown that, in large markets, K is asymptotically equivalent to N times the quantity defined in (5) below) and heuristic methods (which often exploited known order statistics results) were used to compute the order of magnitude of the error term. The important problem of a higher order distributional approximation (of a central limit theorem type) to (K, M) has remained so far open.

In this paper, we solve that problem using a strong approximation approach, establishing a normal approximation to the quantities of interest, together with a bound for the approximation rate. We note that our approach readily extends to the analysis of revenue maximisation mechanisms and constrained optimal mechanisms, through appropriate transformations of the functions F and G (for the problem formulation, see e.g. [13]).

Now we turn to stating additional assumptions on the model that we will need. First of all, in our large market setup, the d.f.'s F and G remain fixed and satisfy the following standard mechanism design assumption:

(A1) *The d.f.'s F and G are absolutely continuous, with respective densities $f(x)$ and $g(x)$ bounded and bounded away from zero on their respective supports $[a, b]$ and $[c, d]$, such that $(a, b) \cap (c, d) \neq \emptyset$ (in other words, $a < d$ and $c < b$).*

It is natural and convenient to consider, for our fixed d.f.'s F and G , a net of markets

$$\mathcal{M}_\alpha := \langle \alpha := (M, N), F, G \rangle, \quad \alpha \in \mathcal{A},$$

indexed by the directed set \mathcal{A} about which we will make the following assumption:

(A2) *We assume that*

$$\mathcal{A} := \{\alpha = (M, N) \in \mathbb{N}^2 : \lambda_\alpha := MN^{-1} \in I\},$$

where $I := [1 - F(d) + \epsilon, 1/(G(a) + \epsilon)]$ for a fixed $\epsilon > 0$.

Here ϵ is assumed to be small enough so that $I \neq \emptyset$; note that $F(d) > 0$ and $G(a) < 1$ by virtue of the assumption $a < d$, see **(A1)**. Note also that $\inf I > 0$.

The set \mathcal{A} is endowed with the natural preorder: for $\alpha = (N, M)$ and $\alpha' = (N', M')$, one has $\alpha \leq \alpha'$ iff $N \leq N'$ and $M \leq M'$.

We will be interested in the limiting distributions of the Walrasian quantities K_α and Walrasian welfares W_α for the respective markets from the net $\{\mathcal{M}_\alpha\}_{\alpha \in \mathcal{A}}$.

The assumption $\lambda_\alpha \in I$ from **(A2)** is natural for economic applications (see e.g. p. 8 of [12]) and is basically necessary for excluding trivial cases. Indeed, with probability tending to one, for large α the number of V_i 's exceeding d will be equal to $(1 - F(d) + o(1))N$. So if $\lambda_\alpha < 1 - F(d) - \epsilon$ for a fixed $\epsilon > 0$, it would mean that the total number of sellers $M = \lambda_\alpha N$ would be less than the number of buyers with valuations higher than the maximum possible productions cost d , meaning that all sellers trade and there is rationing on the demand side of the market. Likewise, the

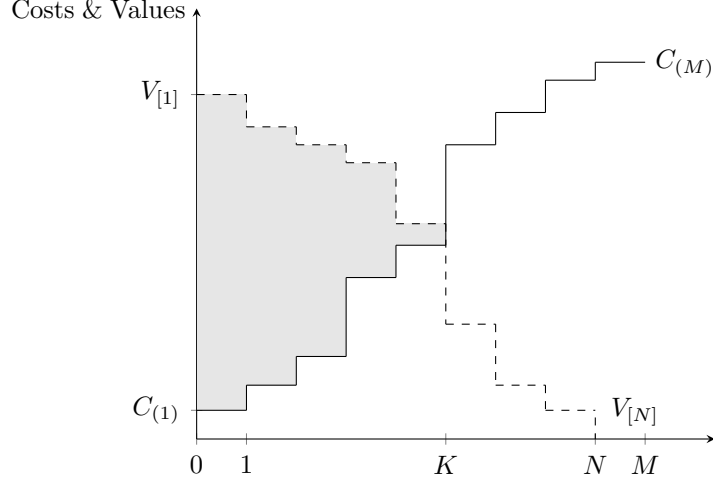


Figure 1: The Walrasian quantity K is given by the abscissa of the intersection of the plots of the buyer and seller order statistics, and the value of Walrasian welfare W is equal to the area of the shaded region.

situation when $\lambda_\alpha > 1/G(a)$ corresponds to a market with excess supply (all buyers trade and there is rationing on the supply side of the market).

To state the main results, we need some further notation. We will frequently deal with variable scaled functions of the form $h(\lambda_\alpha^{-1}t)$. For convenience, for $h : [0, 1] \rightarrow \mathbb{R}$ we define

$$\widehat{h}(s) := h(s \wedge 1), \quad s \geq 0.$$

Note that the function $\widehat{h}(\lambda_\alpha^{-1}t)$, $t \in [0, 1]$, is well-defined for all $\lambda_\alpha > 0$.

Using notation $h^{(-1)}$ for the inverse of function h (to avoid confusion with the reciprocal h^{-1}), introduce the functions

$$E_\alpha(t) := F^{(-1)}(1 - t) - \widehat{G^{(-1)}}(\lambda_\alpha^{-1}t), \quad t \in [0, 1], \quad (3)$$

and put

$$\mathcal{H}(h) := \sup\{t \in (0, 1) : h(t) \geq 0\} \quad (4)$$

(which is well-defined for any $h : [0, 1] \rightarrow \mathbb{R}$ with $h(0) > 0$) and

$$t_\alpha := \mathcal{H}(E_\alpha) \in (0, \lambda_\alpha \wedge 1), \quad (5)$$

where the right relation holds due to the assumption on I from **(A2)**.

Due to the a.s. convergence of empirical quantile functions (e.q.f.'s) to the theoretical ones, for large markets the function $E_\alpha(t/N)$ approximates the difference between the step-functions whose plots are depicted, respectively, by the dashed and solid lines in Figure 1 (cf. (12) below). So $t_\alpha N$ will be the “first order approximation” to K_α , while the integral of E_α over $(0, t_\alpha)$ will specify such deterministic

approximation to W_α . The “second order approximation” to both K_α and W_α will be obtained in this paper using the second order approximation to e.q.f.’s provided by the sum of the theoretical quantile function and a Brownian bridge process.

Now observe that, provided that f, g are continuous inside their respective supports, the function $E_\alpha(t)$ is continuously differentiable for $t \in (0, 1) \setminus \{\lambda_\alpha\}$, and

$$E'_\alpha(t) = -\frac{1}{f(F^{(-1)}(1-t))} - \frac{1}{\lambda_\alpha g(G^{(-1)}(\lambda_\alpha^{-1}t))}, \quad t \in (0, \lambda_\alpha \wedge 1). \quad (6)$$

We will need one more technical assumption on the d.f.’s F and G :

(A3) *The densities f and g are differentiable on (a, b) and (c, d) , respectively. Moreover, the functions*

$$\frac{d}{dt} \frac{1}{f(F^{(-1)}(t))} = -\frac{f'(F^{(-1)}(t))}{f^3(F^{(-1)}(t))} \quad \text{and} \quad \frac{d}{dt} \frac{1}{g(G^{(-1)}(t))} = -\frac{g'(G^{(-1)}(t))}{g^3(G^{(-1)}(t))}$$

are bounded on $(0, 1)$.

Finally, let

$$\sigma_\alpha^2 := \frac{1}{(E'_\alpha(t_\alpha))^2} \left[\frac{t_\alpha(1-t_\alpha)}{f^2(F^{(-1)}(1-t_\alpha))} + \frac{t_\alpha(1-\lambda_\alpha^{-1}t_\alpha)}{\lambda_\alpha^2 g^2(G^{(-1)}(\lambda_\alpha^{-1}t_\alpha))} \right], \quad (7)$$

and

$$\varsigma_\alpha^2 := 2 \int_0^{t_\alpha} S_\alpha(t) dt, \quad (8)$$

where the function $S_\alpha(t)$, $t \in (0, t_\alpha)$, is given by

$$\frac{1-t}{f(F^{(-1)}(1-t))} \int_{F^{(-1)}(1-t)}^b (1-F(x)) dx + \frac{1-\lambda_\alpha^{-1}t}{g(G^{(-1)}(\lambda_\alpha^{-1}t))} \int_c^{G^{(-1)}(\lambda_\alpha^{-1}t)} G(x) dx.$$

Our main result is the following strong approximation theorem. It implies that, under the above assumptions, both the Walrasian quantity and Walrasian welfare are asymptotically normal for large markets with approximate distributions $N(Nt_\alpha, N\sigma_\alpha^2)$ and $N(N \int_0^{t_\alpha} E_\alpha(t) dt, N\varsigma_\alpha^2)$, respectively. Moreover, their joint distribution is also asymptotically normal. We also give upper bounds for the convergence rates.

Theorem 1. *Under assumptions (A1)–(A3), for the net of markets $\{\mathcal{M}_\alpha\}_{\alpha \in \mathcal{A}}$ there exist a net $\{\mathcal{R}_\alpha\}_{\alpha \in \mathcal{A}}$ of realizations of these markets on a common probability space together with a net of bivariate normal random vectors $\{(Z_\alpha^{(1)}, Z_\alpha^{(2)})\}_{\alpha \in \mathcal{A}}$ with zero means and*

$$\text{Var}(Z_\alpha^{(1)}) = \sigma_\alpha^2, \quad \text{Var}(Z_\alpha^{(2)}) = \varsigma_\alpha^2, \quad \text{Cov}(Z_\alpha^{(1)}, Z_\alpha^{(2)}) = \varkappa_\alpha := -S_\alpha(t_\alpha)/E'_\alpha(t_\alpha),$$

such that

$$\limsup_{\alpha} \frac{N^{1/4}}{(\ln N)^{1/2}} \left| \frac{K_{\alpha} - Nt_{\alpha}}{N^{1/2}} - Z_{\alpha}^{(1)} \right| < \infty \quad a.s. \quad (9)$$

and

$$\limsup_{\alpha} \frac{N^{1/2}}{\ln N} \left| \frac{1}{N^{1/2}} \left(W_{\alpha} - N \int_0^{t_{\alpha}} E_{\alpha}(t) dt \right) - Z_{\alpha}^{(2)} \right| < \infty \quad a.s.$$

As an illustration to the assertion of the theorem, Figure 2 shows the scatterplot of 10^4 independent realizations of the centered and scaled (as per the statement of Theorem 1) random vector (K_{α}, W_{α}) in the case when $\alpha = (250, 500)$, $F(x) = x$ and $G(x) = x^2$ for $x \in [0, 1]$, together with the 0.25, 0.5 and 0.95 ellipsoidal quantiles for the sample. Figure 3 displays these empirical ellipsoidal quantiles together with the respective (theoretical) ellipsoidal quantiles for the approximating distribution of $(Z_{\alpha}^{(1)}, Z_{\alpha}^{(2)})$. Both plots were generated using MATHEMATICA 10.

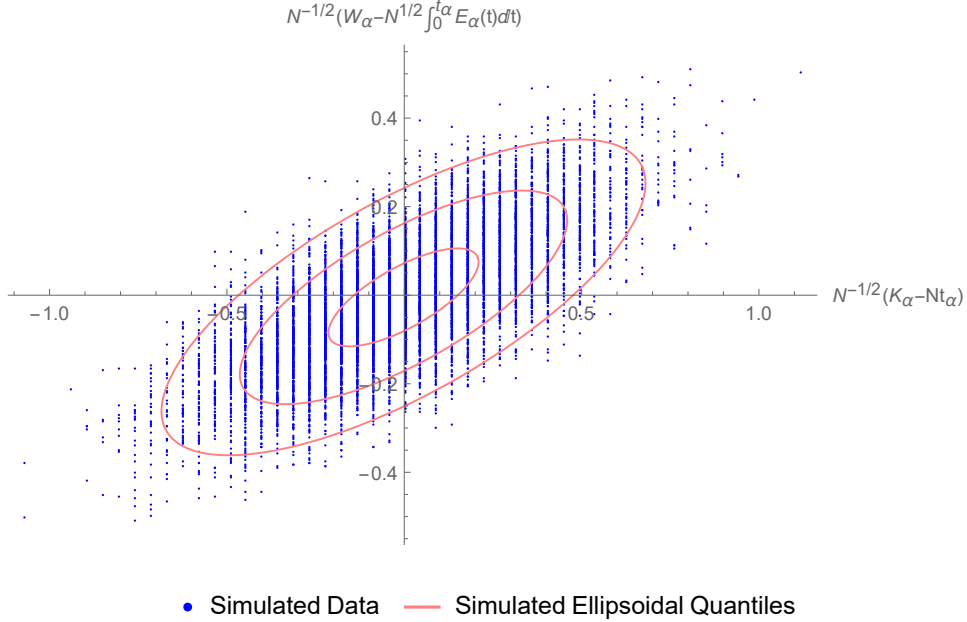


Figure 2: The scatterplot of simulated 10^4 independent scaled copies of (K_{α}, W_{α}) and its ellipsoidal quantiles (see the paragraph following Theorem 1 for more detail).

Remark 1. Note that $N \asymp M$ under assumption **(A2)**, so one could state the assertions of the theorem in a similar way using M rather than N as well. Note also that $\varkappa_{\alpha} > 0$ (as one would expect, of course) since $S_{\alpha}(t_{\alpha}) > 0$ and $E'_{\alpha}(t_{\alpha}) < 0$.

Remark 2. The true rate of convergence of the distribution of K_{α} to the normal one is most likely $N^{-1/2}$, as indicated by the bounds (30) below established in the special case when $F = G$. So the crudeness of the bound in (9) seems to be due to the method of proof employed.

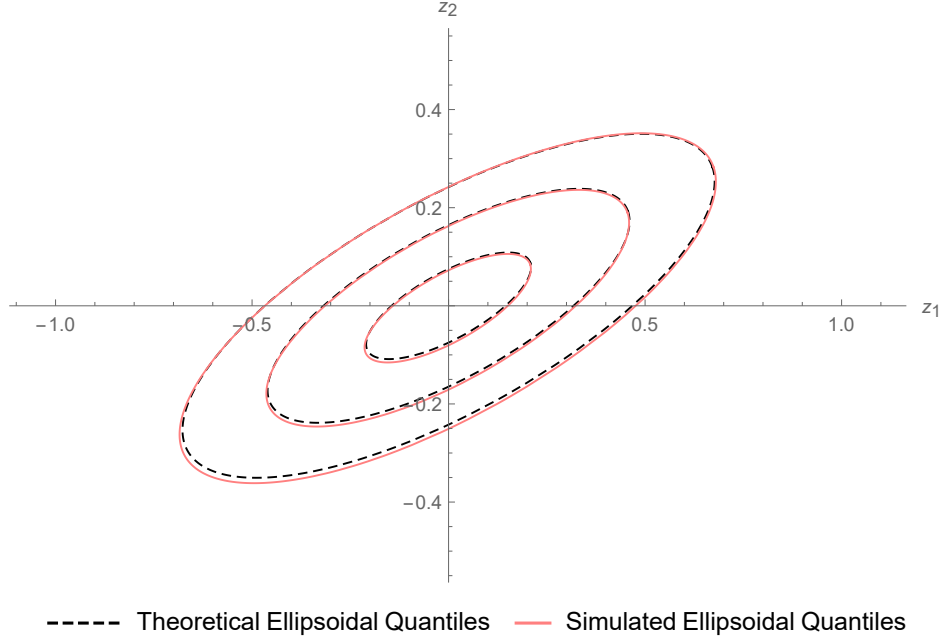


Figure 3: The empirical ellipsoidal quantiles from Figure 2 together with the ellipsoidal quantiles of the approximating bivariate normal distribution.

The rest of the paper is structured as follows. The proof of Theorem 1 is given in Section 2. The special case when $F = G$ is further discussed in Section 3.

2 Proof of the main result

The proof will use the common probability space method and transformed uniform e.q.f.'s, so we will begin with introducing basic notation and recalling some key facts.

For a sample X_1, \dots, X_n with order statistics $X_{(1)} \leq \dots \leq X_{(n)}$, the respective e.q.f. $\mathbb{X}_n(t)$, $t \in [0, 1]$, is defined by

$$\mathbb{X}_n(0) := X_{(1)}, \quad \mathbb{X}_n(t) := X_{(i)}, \quad (i-1)n^{-1} < t \leq in^{-1}, \quad i = 1, \dots, n,$$

For the e.q.f. $\mathbb{U}_n(t)$ constructed from a sample of $U(0, 1)$ -distributed i.i.d. U_1, \dots, U_n ,

$$\mathbb{R}_n(t) := \sqrt{n}(\mathbb{U}_n(t) - t), \quad t \in [0, 1],$$

denotes the respective uniform quantile process. Recall that both \mathbb{U}_n and \mathbb{R}_n are random elements of the Skorokhod space $\mathcal{D}[0, 1]$ of càdlàg functions on $[0, 1]$.

By the well-known Donsker's theorem (see e.g. Section 14 of [1]), as $n \rightarrow \infty$, the distribution of \mathbb{R}_n in $\mathcal{D}[0, 1]$ converges weakly to that of the Brownian bridge process \mathbb{B} . A sharp bound on the rate of convergence is given in [5]. A corollary of that result states that there exists a probability space carrying a sequence of processes $\mathbb{R}_n^* \stackrel{d}{=} \mathbb{R}_n$, $n = 1, 2, \dots$, and a Brownian bridge process \mathbb{B} such that, as

$n \rightarrow \infty$,

$$\|\mathbb{R}_n^* - \mathbb{B}\|_\infty = O(n^{-1/2} \ln n) \quad \text{a.s.},$$

where $\|h\|_\infty := \sup_{t \in [0,1]} |h(t)|$, $h \in \mathcal{D}[0,1]$.

Therefore, as $n \rightarrow \infty$,

$$\mathbb{U}_n(t) = t + n^{-1/2} \mathbb{R}_n(t) \stackrel{d}{=} t + n^{-1/2} \mathbb{R}_n^*(t) = t + n^{-1/2} \mathbb{B}(t) + \theta_n(t) n^{-1} \ln n \quad (10)$$

for $t \in [0,1]$, where $\|\theta_n\|_\infty = O(1)$ a.s.

Proof of Theorem 1. It is easy to see from (1), (4) and (5) that

$$\delta_\alpha := K_\alpha N^{-1} - t_\alpha = \mathcal{H}(\mathbb{E}_\alpha) - \mathcal{H}(E_\alpha), \quad (11)$$

where

$$\mathbb{E}_\alpha(t) := \mathbb{V}_N(1-t) - \widehat{\mathbb{C}}_M(\lambda_\alpha^{-1}t), \quad t \in [0,1], \quad (12)$$

and \mathbb{V}_N and \mathbb{C}_M denote the e.q.f.'s for the samples of buyer valuations V_1, \dots, V_N and seller costs C_1, \dots, C_M , respectively. Now we will analyze the behavior of δ_α for “large” α when market realizations \mathcal{R}_α are defined using the common probability space construction (10).

To that end, we will assume that

$$V_i = V_{i,N} := F^{(-1)}(U_{i,N}^V) \quad \text{and} \quad C_j = C_{j,M} := G^{(-1)}(U_{j,M}^C), \quad (13)$$

where $\{U_{1,N}^V, \dots, U_{N,N}^V\}_{N \geq 1}$ and $\{U_{1,M}^C, \dots, U_{M,M}^C\}_{M \geq 1}$ are independent triangular arrays of row-wise independent $U(0,1)$ -distributed random variables, which are defined on a common probability space with Brownian bridges \mathbb{B}^V and \mathbb{B}^C that are independent of each other so that, as $N, M \rightarrow \infty$, for the respective e.q.f.'s one has

$$\begin{aligned} \mathbb{U}_N^V(t) &= t + N^{-1/2} \mathbb{B}^V(t) + \theta_N^V(t) N^{-1} \ln N, \quad t \in [0,1], \\ \mathbb{U}_M^C(t) &= t + M^{-1/2} \mathbb{B}^C(t) + \theta_M^C(t) M^{-1} \ln M, \quad t \in [0,1], \end{aligned} \quad (14)$$

where $\|\theta_N^V\|_\infty = O(1)$ and $\|\theta_M^C\|_\infty = O(1)$ a.s.

In this construction, the e.q.f.'s \mathbb{V}_N and \mathbb{C}_M in (12) are the ones for the samples in (13), i.e. are given by the compositions

$$\mathbb{V}_N(t) := (F^{(-1)} \circ \mathbb{U}_N^V)(t) \quad \text{and} \quad \mathbb{C}_M(t) := (G^{(-1)} \circ \mathbb{U}_M^C)(t). \quad (15)$$

So to analyse the asymptotic distribution of K_α , we will now turn to the asymptotic behaviour of \mathbb{E}_α specified in (12) where the terms on the right-hand side are given by (15).

Using (15), (14) and conditions **(A1)**, **(A3)** to take Taylor series expansions with two terms and Lagrange's form of the remainder gives, after elementary transformations,

$$\begin{aligned} \mathbb{V}_N(1-t) &= F^{(-1)}(1-t + N^{-1/2} \mathbb{B}^V(1-t) + \theta_N^V(1-t) N^{-1} \ln N) \\ &= F^{(-1)}(1-t) + \frac{N^{-1/2} \mathbb{B}^V(1-t) + \vartheta_N^V(t) N^{-1} \ln N}{f(F^{(-1)}(1-t))}, \quad t \in (0,1), \end{aligned} \quad (16)$$

and, for $t \in (0, 1 \wedge \lambda_\alpha)$,

$$\begin{aligned}\mathbb{C}_M(\lambda_\alpha^{-1}t) &= G^{(-1)}(\lambda_\alpha^{-1}t + M^{-1/2}\mathbb{B}^C(\lambda_\alpha^{-1}t) + \theta_M^C(\lambda_\alpha^{-1}t)M^{-1}\ln M) \\ &= G^{(-1)}(\lambda_\alpha^{-1}t) + \frac{M^{-1/2}\mathbb{B}^C(\lambda_\alpha^{-1}t) + \vartheta_M^C(t)M^{-1}\ln M}{g(G^{(-1)}(\lambda_\alpha^{-1}t))},\end{aligned}\quad (17)$$

where $\|\vartheta_M^C\|_\infty + \|\vartheta_N^V\|_\infty = O(1)$ a.s.

Substituting (16) and (17) into the representation for \mathbb{E}_α in (12) and using (3) gives

$$\mathbb{E}_\alpha(t) = E_\alpha(t) + N^{-1/2}\mathbb{Z}_\alpha(t) + \varphi_\alpha(t)N^{-1}\ln N, \quad (18)$$

where $\|\varphi_\alpha\| = O(1)$ a.s. and

$$\mathbb{Z}_\alpha(t) := \frac{\mathbb{B}^V(1-t)}{f(F^{(-1)}(1-t))} - \frac{\widehat{\mathbb{B}^C}(\lambda_\alpha^{-1}t)}{\lambda_\alpha^{1/2}g(\widehat{G^{(-1)}}(\lambda_\alpha^{-1}t))}, \quad t \in (0, 1). \quad (19)$$

One can see from (11) and (18) that if the functional \mathcal{H} were differentiable in a suitable sense at the “point” E_α , one could derive the desired asymptotic normality of K_α using a suitable version of the delta method. So we will now turn to studying the local properties of \mathcal{H} at E_α .

Since f , g and λ_α are bounded, we see from (6) that

$$-\gamma := \sup_{\alpha \in \mathcal{A}} \sup_{t \in (0, \lambda_\alpha \wedge 1)} E'_\alpha(t) < 0.$$

As $E_\alpha(t_\alpha) = 0$, the above implies that

$$E_\alpha(t) \geq \gamma(t_\alpha - t), \quad 0 \leq t < t_\alpha; \quad E_\alpha(t) \leq \gamma(t_\alpha - t), \quad t_\alpha < t \leq \lambda_\alpha \wedge 1.$$

Hence, for any $v \in \mathcal{D}[0, 1]$, setting $\eta := \|v\|_\infty$, one has

$$\begin{aligned}\inf_{0 \leq t < t_\alpha - \eta/\gamma} (E_\alpha(t_\alpha) + v(t)) &> \gamma\eta/\gamma - \eta = 0, \\ \sup_{t_\alpha + \eta/\gamma < t \leq \lambda_\alpha \wedge 1} (E_\alpha(t) + v(t)) &< -\gamma\eta/\gamma + \eta = 0,\end{aligned}$$

so that the plot of $E_\alpha(t_\alpha) + v(t)$ must “dive” under zero within η/γ of t_α :

$$|\mathcal{H}(E_\alpha + v) - \mathcal{H}(E_\alpha)| \equiv |\mathcal{H}(E_\alpha + v) - t_\alpha| \leq \eta/\gamma = \|v\|_\infty/\gamma.$$

Thus, the functional \mathcal{H} is Lipschitz continuous in the uniform topology at the point E_α . Using this, (11) and (18), we have

$$\delta_\alpha = O(N^{-1/2}) \quad \text{a.s.} \quad (20)$$

Further, \mathbb{E}_α is right continuous, it follows from the definition of \mathcal{H} that

$$\mathbb{E}_\alpha(t_\alpha + \delta_\alpha) \leq 0. \quad (21)$$

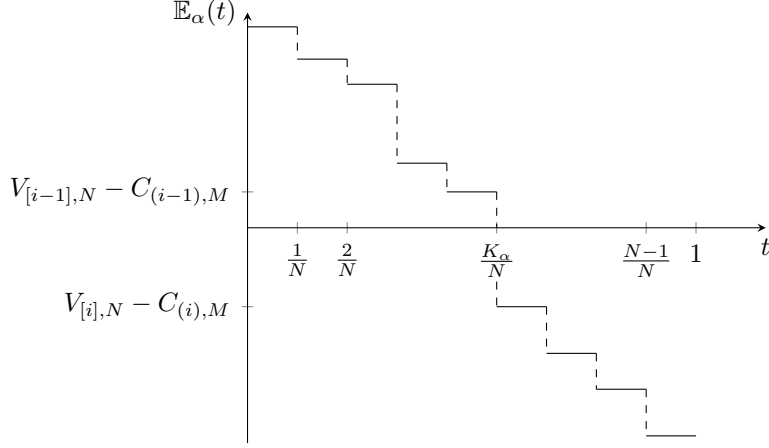


Figure 4: A graphical illustration of inequalities (21), (22). Recall that $K_\alpha N^{-1} = \mathcal{H}(\mathbb{E}_\alpha(t)) = t_\alpha + \delta_\alpha$.

On the other hand, using the order statistics $V_{[i-1],N}$ (descending) and $C_{(i),M}$ (ascending) for our samples (13) (cf. the definitions just before (1)), we obtain (see Fig. 4 for an illustration of the first inequality)

$$\begin{aligned} \mathbb{E}_\alpha(t_\alpha + \delta_\alpha) &\geq -\max_{2 \leq i \leq K_\alpha} [(V_{[i-1],N} - V_{[i],N}) + (C_{(i),M} - C_{(i-1),M})] \\ &\geq -\max_{2 \leq i \leq N} (V_{[i-1],N} - V_{[i],N}) - \max_{2 \leq i \leq M} (C_{(i),M} - C_{(i-1),M}). \end{aligned} \quad (22)$$

From a well-known result regarding maximal uniform spacings (see e.g. [8]) and representation (13), we have

$$\max_{2 \leq i \leq N} (V_{[i-1],N} - V_{[i],N}) \leq \frac{\max_{2 \leq i \leq N} (U_{[i-1],N}^V - U_{[i],N}^V)}{\min_{t \in [a,b]} f(t)} = O(N^{-1} \ln N) \quad \text{a.s.}$$

Since the second term on the right-hand side of (22) has the same order of magnitude, (21) and (22) now yield

$$\mathbb{E}_\alpha(t_\alpha + \delta_\alpha) = O(N^{-1} \ln N) \quad \text{a.s.}$$

Therefore (18) implies that

$$E_\alpha(t_\alpha + \delta_\alpha) + N^{-1/2} \mathbb{Z}_\alpha(t_\alpha + \delta_\alpha) = O(N^{-1} \ln N). \quad (23)$$

Under assumption **(A3)** we also have

$$E_\alpha(t_\alpha + \delta_\alpha) = E_\alpha(t_\alpha) + \delta_\alpha E'_\alpha(t_\alpha) + O(\delta_\alpha^2) = \delta_\alpha E'_\alpha(t_\alpha) + O(\delta_\alpha^2) \quad \text{a.s.} \quad (24)$$

Combining (23) with (24) and using (20) gives

$$\begin{aligned} \delta_\alpha E'_\alpha(t_\alpha) &= -N^{-1/2} \mathbb{Z}_\alpha(t_\alpha + \delta_\alpha) + O(N^{-1} \ln N) \\ &= -N^{-1/2} \mathbb{Z}_\alpha(t_\alpha) + N^{-1/2} \psi_\alpha \omega_{\mathbb{Z}_\alpha}(\delta_\alpha) + O(N^{-1} \ln N) \end{aligned} \quad (25)$$

with $|\psi_\alpha| \leq 1$, where $\omega_h(\cdot)$ stands for the modulus of continuity of the function h on $[0, 1]$. Recall that, for a Brownian bridge process \mathbb{B} , one has

$$\limsup_{\delta \downarrow 0} \frac{w_{\mathbb{B}}(\delta)}{\sqrt{2\delta \ln(1/\delta)}} = 1 \quad \text{a.s.}$$

(see e.g. Theorem 1.4.1 in [7]). As this holds for both \mathbb{B}^V and \mathbb{B}^C , and λ_α , f and g are bounded away from zero, it follows from (19) that

$$\omega_{\mathbb{Z}_\alpha}(\delta_\alpha) = O(\delta_\alpha^{1/2} \ln^{1/2}(1/\delta_\alpha)) = O(N^{-1/4}(\ln N)^{1/2}) \quad \text{a.s.}$$

Hence (25) yields now

$$\delta_\alpha = -\frac{\mathbb{Z}_\alpha(t_\alpha)}{N^{1/2}E'_\alpha(t_\alpha)} + O(N^{-3/4}(\ln N)^{1/2}) \quad \text{a.s.} \quad (26)$$

Since the processes \mathbb{B}^V and \mathbb{B}^C in (19) are independent and $\mathbb{B}(t) \sim N(0, t(1-t))$, we immediately see that $\mathbb{Z}_\alpha(t_\alpha)$ is normally distributed with zero mean and variance

$$\frac{t_\alpha(1-t_\alpha)}{f^2(F^{(-1)}(1-t_\alpha))} + \frac{t_\alpha(1-\lambda_\alpha^{-1}t_\alpha)}{\lambda_\alpha^2 g^2(G^{(-1)}(\lambda_\alpha^{-1}t_\alpha))}.$$

It follows from (7) that

$$Z_\alpha^{(1)} := -\mathbb{Z}_\alpha(t_\alpha)/E'_\alpha(t_\alpha) \sim N(0, \sigma_\alpha^2),$$

and so (26) establishes the first assertion of the theorem.

For the Walrasian welfare, from (2), (11) and (12) we have

$$\begin{aligned} W_\alpha &= \sum_{i=1}^{K_\alpha} (V_{[i],N} - C_{(i),M}) \\ &= N \int_0^{K_\alpha/N} (\mathbb{V}_N(1-t) - \mathbb{C}_M(\lambda_\alpha^{-1}t)) dt = N \int_0^{K_\alpha/N} \mathbb{E}_\alpha(t) dt \\ &= N \int_0^{t_\alpha+\delta_\alpha} (E_\alpha(t) + N^{-1/2}\mathbb{Z}_\alpha(t) + \varphi_\alpha(t)N^{-1} \ln N) dt. \end{aligned} \quad (27)$$

Now note that replacing $\int_0^{t_\alpha+\delta_\alpha}$ with $\int_0^{t_\alpha}$ in the last line will only introduce an error $O(1)$ a.s. Indeed, since $E_\alpha(t_\alpha) = 0$ and $E'_\alpha(t)$ is bounded in view of our assumptions **(A1)** and **(A2)** (cf. (6)), one has from (20) that

$$N \int_{t_\alpha}^{t_\alpha+\delta_\alpha} E_\alpha(t) dt = O\left(N \int_{t_\alpha}^{t_\alpha+|\delta_\alpha|} (t-t_\alpha) dt\right) = O(N\delta_\alpha^2) = O(1) \quad \text{a.s.}$$

Further, it is clear from **(A1)**, **(A2)** and (19) that there exists a constant $c < \infty$ such that

$$\max_{t \in [0,1]} |\mathbb{Z}_\alpha(t)| \leq Y := c \left(\max_{t \in [0,1]} |\mathbb{B}^V(t)| + \max_{t \in [0,1]} |\mathbb{B}^C(t)| \right) < \infty \quad \text{a.s.}$$

Hence, again using (20),

$$\left| N^{1/2} \int_{t_\alpha}^{t_\alpha + \delta_\alpha} \mathbb{Z}_\alpha dt \right| \leq N^{1/2} |\delta_\alpha| Y = O(1) \quad \text{a.s.}$$

This proves the desired claim since the contribution of the last term in the integrand in (27) to $\int_{t_\alpha}^{t_\alpha + \delta_\alpha}$ will be even smaller in magnitude (recall that $\|\varphi_\alpha\|_\infty = O(1)$ a.s. and one has (20)).

Thus we obtain from (27) that

$$W_\alpha = N \int_0^{t_\alpha} E_\alpha(t) dt + N^{1/2} \int_0^{t_\alpha} \mathbb{Z}_\alpha(t) dt + O(\ln N) \quad \text{a.s.} \quad (28)$$

It is clear from (19) that \mathbb{Z}_α is a zero mean Gaussian process, and so the second integral on the right-hand side of (28), to be denoted by $Z_\alpha^{(2)}$, is zero mean normal as well. Putting for brevity $f^*(t) := f(F^{(-1)}(t))$, $g^*(t) := g(G^{(-1)}(t))$, recalling the right relation in (5) and also that \mathbb{B}^V and \mathbb{B}^C are independent Brownian bridges, we compute the variance of $Z_\alpha^{(2)}$ as

$$\begin{aligned} \mathbf{E} \left(\int_0^{t_\alpha} \mathbb{Z}_\alpha(t) dt \right)^2 &= \iint_{[0, t_\alpha]^2} \mathbf{E} \mathbb{Z}_\alpha(s) \mathbb{Z}_\alpha(t) ds dt = 2 \iint_{[0, t_\alpha]^2 \cap \{s < t\}} \mathbf{E} \mathbb{Z}_\alpha(s) \mathbb{Z}_\alpha(t) ds dt \\ &= 2 \int_0^{t_\alpha} \int_0^t \left[\frac{\mathbf{E} \mathbb{B}^V(1-s) \mathbb{B}^V(1-t)}{f^*(1-s) f^*(1-t)} + \frac{\mathbf{E} \mathbb{B}^C(\lambda_\alpha^{-1}s) \mathbb{B}^C(\lambda_\alpha^{-1}t)}{\lambda_\alpha g^*(\lambda_\alpha^{-1}s) g^*(\lambda_\alpha^{-1}t)} \right] ds dt \\ &= 2 \int_0^{t_\alpha} \int_0^t \left[\frac{s(1-t)}{f^*(1-s) f^*(1-t)} + \frac{s(1-\lambda_\alpha^{-1}t)}{\lambda_\alpha^2 g^*(\lambda_\alpha^{-1}s) g^*(\lambda_\alpha^{-1}t)} \right] ds dt \\ &= 2 \int_0^{t_\alpha} \left[\frac{1-t}{f^*(1-t)} \int_0^t \frac{s ds}{f^*(1-s)} + \frac{1-\lambda_\alpha^{-1}t}{g^*(\lambda_\alpha^{-1}t)} \int_0^t \frac{s ds}{\lambda_\alpha^2 g^*(\lambda_\alpha^{-1}s)} \right] dt. \end{aligned} \quad (29)$$

Changing variables $x := F^{(-1)}(1-s)$ and $x := G^{(-1)}(\lambda_\alpha^{-1}s)$, respectively, in the two integrals inside the square brackets in the last line in (29), we obtain the desired representation (8) for ζ_α^2 . Now the second assertion of the theorem follows from (28).

To complete the proof of the theorem, it remains to compute the covariance

$$\mathbf{E} Z_\alpha^{(1)} Z_\alpha^{(2)} = -\mathbf{E} \frac{\mathbb{Z}_\alpha(t_\alpha)}{E'_\alpha(t_\alpha)} \int_0^{t_\alpha} \mathbb{Z}_\alpha(t) dt = -\frac{1}{E'_\alpha(t_\alpha)} \int_0^{t_\alpha} \mathbf{E} \mathbb{Z}_\alpha(t_\alpha) \mathbb{Z}_\alpha(t) dt = -\frac{S_\alpha(t_\alpha)}{E'_\alpha(t_\alpha)},$$

where the last equality follows from the calculation of the inner integral in (29) and the definition of the function S_α (following (8)). Theorem 1 is proved. \square

Remark 3. An argument similar to the second part of our proof of asymptotic normality of K_α can be found in the first example in Section 8, Ch. I of [3]. Note also that a direct application of a known general version of the delta method such as the proposition in [4] is not possible in our case as the quantity t_α depends on α and does not converge to any value.

Remark 4. The reader is referred to [6] for the minimal boundedness conditions that must be imposed on f' and g' for (17) and (16) to hold. It turns out that, for the approximation of K_α , these boundedness conditions may be weakened further as they are only required in a neighbourhood of t_α .

3 The special case $F = G$

In the interesting special case $F = G$, the distribution of K_α is clearly independent of F . Moreover, it is known to be hypergeometric (see the proof of part (ii) below for a more precise statement) and the parameters of the approximating normal law for K_α admit simple explicit representations as functions of α . In addition, in that case one can establish a better convergence rate to the normal distribution than the one claimed in Theorem 1.

The distribution of W_α (and hence that of (K_α, W_α)) and the approximating normal law clearly do depend on F . However, one can also obtain simple closed formulae for the parameters of the approximating normal laws when the common distribution $F = G$ is uniform on (a, b) . We will only deal here with the case $a = 0$, $b = 1$, the results in the general case following in a straightforward way, and state our results in the form of the following theorem.

Theorem 2. *If $F = G$ then:*

(i) *one has*

$$t_\alpha = \frac{\lambda_\alpha}{1 + \lambda_\alpha}, \quad \sigma_\alpha^2 = \frac{\lambda_\alpha^2}{(1 + \lambda_\alpha)^3};$$

(ii) *there exist constants $C_1, C_2 \in (0, \infty)$ such that*

$$\frac{C_1}{N^{1/2}} \leq \sup_x \left| \mathbf{P} \left(\frac{K_\alpha - Nt_\alpha}{\sigma_\alpha N^{1/2}} \leq x \right) - \Phi(x) \right| \leq \frac{C_2}{N^{1/2}}, \quad (30)$$

where Φ is the standard normal d.f.

(iii) *If, moreover, $F(t) = G(t) = t$ for $t \in (0, 1)$, then*

$$\int_0^{t_\alpha} E_\alpha(t) dt = \frac{\lambda_\alpha}{2(1 + \lambda_\alpha)}, \quad \varsigma_\alpha^2 = \frac{\lambda_\alpha(1 + 3\lambda_\alpha + \lambda_\alpha^2)}{12(1 + \lambda_\alpha)^3}, \quad \varkappa_\alpha = \frac{\lambda_\alpha^2}{2(1 + \lambda_\alpha)^3}. \quad (31)$$

Remark 5. Observe that, in the case from part (iii), the correlation coefficient between the components of the approximating normal distribution can be easily found to be equal to $\sqrt{3/(\lambda_\alpha^{-1} + 3 + \lambda_\alpha)}$. That quantity attains its maximum value $\sqrt{3/5}$ at $\lambda_\alpha = 1$ (i.e. when $M = N$) and vanishes as $\lambda_\alpha \vee \lambda_\alpha^{-1} \rightarrow \infty$.

Proof. (i) When $G = F$, the equation $E_\alpha(t_\alpha) = 0$ for t_α (see (3), (5)) turns into $F^{(-1)}(\lambda_\alpha^{-1}t_\alpha) = F^{(-1)}(1 - t_\alpha)$, which means that $\lambda_\alpha^{-1}t_\alpha = 1 - t_\alpha$, yielding the desired representation for t_α .

Next, in this case for $t = t_\alpha$ the value of (6) turns into

$$E'_\alpha(t_\alpha) = -\frac{1}{t_\alpha f(F^{(-1)}(1 - t_\alpha))},$$

so that (7) becomes

$$\sigma_\alpha^2 = t_\alpha^2 [t_\alpha(1 - t_\alpha) + t_\alpha^2 \lambda_\alpha^{-2}] = t_\alpha^2(1 - t_\alpha^2) = \frac{\lambda_\alpha^2}{(1 + \lambda_\alpha)^3}.$$

(ii) As we already pointed out, in the special case when $F = G$ the exact distribution of K_α does not depend on F (which is obvious from the definition of K_α). In fact, as shown in [13], K_α has the hypergeometric distribution $\text{Hg}(N, M, N + M)$. It is well known that, as $M + N \rightarrow \infty$ and the quantity λ_α remains bounded away from 0 and 1 (which is ensured by our assumption **(A2)**), that distribution can be approximated by a normal one (see e.g. Theorem 2.1 in [11]). The stated bounds (30) follow from the results established in Theorem 2.2 in [11].

(iii) First note that, in the case of the uniform on $(0, 1)$ distributions $F = G$, one has $F(-1)(t) \equiv G(-1)(t) \equiv t$ on $(0, 1)$, so that

$$\int_0^{t_\alpha} E_\alpha(t) dt = \int_0^{t_\alpha} (1 - t - t/\lambda_\alpha) dt = \int_0^{t_\alpha} (1 - t/t_\alpha) dt = \frac{t_\alpha}{2} = \frac{\lambda_\alpha}{2(1 + \lambda_\alpha)}.$$

Next, since $f(t) \equiv g(t) \equiv 1$ on $(0, 1)$, we also have

$$\begin{aligned} S_\alpha(t) &= (1 - t) \int_{1-t}^1 (1 - x) dx + (1 - \lambda_\alpha^{-1}t) \int_0^{\lambda_\alpha^{-1}t} x dx \\ &= \frac{1}{2}t^2(1 - t) + \frac{1}{2} \left(\frac{t}{\lambda_\alpha} \right)^2 \left(1 - \frac{t}{\lambda_\alpha} \right), \quad t \in (0, t_\alpha). \end{aligned}$$

Integrating this expression in t from 0 to t_α yields the second formula in (31). To get the last formula in (31), we just note that $E_\alpha(t_\alpha)' = -1/t_\alpha$ and, as $t_\alpha/\lambda_\alpha = 1 - t_\alpha$, one has $S_\alpha(t_\alpha) = \frac{1}{2}t_\alpha^2(1 - t_\alpha) + \frac{1}{2}(1 - t_\alpha)^2t_\alpha = \frac{1}{2}t_\alpha(1 - t_\alpha)$, and so $\varkappa_\alpha = -S_\alpha(t_\alpha)/E_\alpha(t_\alpha)' = \frac{1}{2}t_\alpha^2(1 - t_\alpha)$. □

Acknowledgements

E. Muir's research was supported by the Elizabeth and Vernon Puzey Scholarship and the Australian Research Council through the Laureate Fellowship FL130100039 and the ARC Centre of Excellence for Mathematical and Statistical Frontiers of Big Data, Big Models, New Insights. K. Borovkov's work was supported by the ARC grant DP150102758.

References

- [1] BILLINGSLEY, P. P. (1999). *Convergence of Probability Measures*, 2nd edn. Wiley, New York.
- [2] BÖRGERS, T. (2015). *An Introduction to the Theory of Mechanism Design*. Oxford University Press, New York.
- [3] BOROVKOV, A. A. (1998). *Mathematical Statistics*. Gordon & Breach, Amsterdam.

- [4] BOROVKOV, K. A. (1985). A note on a limit theorem for differentiable mappings. *Ann. Prob.* **13**, 1018–1021.
- [5] CSÖRGŐ, M. AND RÉVÉSZ, P. (1975). Some notes on the empirical distribution function and the quantile process. In *Limit Theorems of Probability Theory*, vol. 11. North Holland, Amsterdam, pp. 59–71.
- [6] CSÖRGŐ, M. AND RÉVÉSZ, P. (1978). Strong approximations of the quantile process. *Ann. Statist.* **6**, 882–894.
- [7] CSÖRGŐ, M. AND RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York.
- [8] DEVROYE, L. (1982). A log log law for maximal uniform spacings. *Ann. Prob.* **10**, 863–868.
- [9] GRESIK, T. A. AND STRATTERTHWAITE, M. A. (1989). The rate at which a simple market converges to efficiency as the number of traders increases: An asymptotic result for optimal trading mechanisms. *J. Economic Theor.* **48**, 304–332.
- [10] KRISHNA, V. (2009). *Auction Theory*, 2nd edn. Academic Press.
- [11] LAHIRI, S. N. AND CHATTERJEE, A. (2007). A Berry-Esseen theorem for hypergeometric probabilities under minimal conditions. *Proc. Amer. Math. Soc.* **135**, 1535–1545.
- [12] LAUERMANN, S. (2013). Dynamic matching and bargaining games: A general approach. *American Economic Review*. **103**, 663–689.
- [13] LOERTSCHER, S., MUIR, E. V. AND TAYLOR, P. G. (2015). A general non-central hypergeometric distribution. To appear in *Commun. Statist. – Theory and Methods*.
- [14] MCAFEE, R. P. (1992). A dominant strategy double auction. *J. of Economic Theor.* **56**, 434–450.
- [15] MUIR, E. V. (2013). Mechanism design and the $\langle n, m \rangle$ trade problem. Master’s thesis, University of Melbourne School of Mathematics and Statistics. <http://www.ms.unimelb.edu.au/documents/thesis/EllenVictoriaMuir.pdf>
- [16] WILLIAMS, S. R. (1999). A characterization of efficient, Bayesian incentive compatible mechanisms. *Economic Theory*. **14**, 155–180.